NON-PARAMETRIC ESTIMATION OF LOCATION

by

M. V. JOHNS, JR.

TECHNICAL REPORT NO. 182

NOVEMBER 2, 1971

•

DEPARTMENT OF STATISTICS

STANFORD   UNIVERSITY

STANFORD, CALIFORNIA

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*	2A. REPORT SECURITY CLASSIFICATION
DEPARTMENT OF STATISTICS	
STANFORD UNIVERSITY CALIFORNIA	2B. GROUP

3. REPORT TITLE

NON-PARAMETRIC ESTIMATION OF LOCATION

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

TECHNICAL REPORT

5. AUTHOR(S) *(Last name, first name, initial)*

JOHNS, JR. M. V.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
November 2, 1971	29	14

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
N00014-67-A-0112-0053	
b. PROJECT NO.	#182
NR-042-267	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	#41 NSF 27550

10. AVAILABILITY/LIMITATION NOTICES

Distribution of document is unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Logistics & Mathematical Sciences Branch Office of Naval Research, Arlington, VA.

13. ABSTRACT

A sequence of asymptotically normally distributed estimators of location is presented, having the property that, for any $\epsilon > 0$, all estimators in the sequence beyond an appropriate point have asymptotic variances within ϵ of the Cramer-Rao lower bound, uniformly for all symmetric distributions in a non-parametric family constrained only by regularity conditions. The simplest non-trivial estimator in this sequence already possesses good efficiency-robustness properties, both asymptotically and for small sample sizes. This estimator is much easier to compute than previously proposed estimators having similar properties, and a good non-parametric estimate of the variance of the location estimator is produced as a byproduct.

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
non-parametric						
estimator						
location						
efficiency-robust						

INSTRUCTIONS

1. **ORIGINATING ACTIVITY:** Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (*corporate author*) issuing the report.

2a. **REPORT SECURITY CLASSIFICATION:** Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. **GROUP:** Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. **REPORT TITLE:** Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. **DESCRIPTIVE NOTES:** If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. **AUTHOR(S):** Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. **REPORT DATE:** Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. **TOTAL NUMBER OF PAGES:** The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. **NUMBER OF REFERENCES:** Enter the total number of references cited in the report.

8a. **CONTRACT OR GRANT NUMBER:** If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. **PROJECT NUMBER:** Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. **ORIGINATOR'S REPORT NUMBER(S):** Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. **OTHER REPORT NUMBER(S):** If the report has been assigned any other report numbers (*either by the originator or by the sponsor*), also enter this number(s).

10. **AVAILABILITY/LIMITATION NOTICES:** Enter any limitations on further dissemination of the report, other than those

imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through
_____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through
_____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through
_____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. **SUPPLEMENTARY NOTES:** Use for additional explanatory notes.

12. **SPONSORING MILITARY ACTIVITY:** Enter the name of the departmental project office or laboratory sponsoring (*paying for*) the research and development. Include address.

13. **ABSTRACT:** Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as (TS), (S), (C), or (U).

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. **KEY WORDS:** Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.

NON-PARAMETRIC ESTIMATION OF LOCATION

by

M. V. Johns, Jr.

1. Introduction.

The problem of "efficiency-robust" estimation of location for symmetric continuous distributions has been treated by a number of authors (as noted in the text below) employing various optimality criteria. The present paper develops a sequence of estimators, indexed by an integer valued parameter k, exhibiting the following rather strong notion of asymptotic "efficiency-robustness": For any k the corresponding estimator is consistent and asymptotically normally distributed (as the sample size n increases) for any F in a large subset $\mathcal{F}$ of the class of symmetric continuous distributions. Furthermore, for any $\epsilon > 0$ the variance of the limiting normal distribution exceeds the Cramér-Rao bound for F by no more than ϵ uniformly for all $F \epsilon \mathcal{F}$, for each sufficiently large k. Thus, for large k the corresponding estimator is (nearly) Best Asymptotically Normal (BAN) for all $F \epsilon \mathcal{F}$.

This concept of optimality is distinguished from some others described below in that the class $\mathcal{F}$ is "non-parametric" and indeed contains the simple parametric families used in other definitions.

Of greater significance is the fact that the simplest non-trivial estimator in the proposed sequence (corresponding to k = 2) exhibits

1

quite high efficiencies for small to moderate sample sizes $(n = 10, 20, 40)$ for a collection of diverse distributions consisting of the normal, the Cauchy, the logistic, the double exponential, and the 10% contaminated normal. These efficiencies relative to the best linear unbiased estimates (BLUE) based on order statistics were obtained by Monte Carlo experiments. The estimator seems slightly less efficient for short-tailed distributions compared to some competitors designed particularly for use with such distributions, but it does considerably better for the long-tailed Cauchy case. The asymptotic efficiency of the estimator may be computed by hand without difficulty.

This estimator is no more difficult to compute than, say, a one sample t-statistic, and the calculation yields as a by-product a good estimate of the variance of the location estimator so that approximate confidence intervals and tests may be constructed. The author believes this computational simplicity to be one of the most important features of the proposed estimator since it enhances the possibility that it may be found useful by statistical practitioners dealing with actual data. Previously proposed procedures exhibiting comparable efficiency-robustness require complex arithmetic manipulations and some involve table lookups or matrix inversions. Considerations of simplicity have previously motivated interest in very easily computable estimators such as trimmed or Windsorized means, symmetric linear combinations of three order statistics, etc. (see, e.g. Gastwirth and Cohen [7]). Such estimators do not exhibit as broad a spectrum of efficiency-robustness as those considered in this paper, nor are they associated with natural

2

non-parametric estimates of their variances.

The various approaches to efficiency-robust estimation may for convenience be classified into two categories: (1) those achieving asymptotic optimality for each member of a specified finite or parametric family of distribution "shapes" (with unspecified scale parameters), and (2) those achieving asymptotic optimality for all members of a non-parametric family of distributions limited only by regularity conditions. Examples of the first category of estimation procedures are given by A. Birnbaum and his associates in [2], [3], [4], J. Gastwirth and H. Rubin [6], R. V. Hogg [8], and V. Miké [10]. In these papers +¹₊e estimators depend explicitly on the parametric collection of distribution shapes chosen and "optimality" means the asymptotic achievement of the Cramér-Rao lower bound for the variance of the estimator, or the maximization of the minimum variance over the specified family of distribution shapes, or admissibility with respect to the family. In all of these cases it is reasonable to expect (and has to some extent been demonstrated) that the proposed estimators will be robust in the sense of performing well for distributions which do not differ too much from the families in terms of which they are defined. A related approach uses the data to select one estimator from a small specified collection of candidates. Such procedures have been proposed by P. Switzer [12] and L. Jaeckel [9].

The approaches falling into the second category mentioned above are inherently more ambitious in their goals since they aim at asymptotic optimality for all distributions in non-parametric families constrained only by regularity conditions. Estimators are regarded

3

as optimal if they achieve the Cramér-Rao bound asymptotically or if they are BAN (or nearly BAN). The possibility of achieving this kind of uniform efficiency was discussed as early as 1956 by C. Stein [11]. Recent proposals for such uniformly efficient estimators have been made by P. K. Bhattacharya [1], C. Van Eeden [14] and K. Takeuchi [13]. The first two of these papers suggest estimators involving pointwise estimates of density functions and are likely to require rather large samples to be effective. Takeuchi's ingenious sequence of estimators (indexed by an integer k) is not only asymptotically optimal for large values of k, but shows good small sample efficiency-robustness (for small values of k). A natural non-parametric estimate of the variance of the estimator is also provided.

The estimators proposed in the present paper fall into the second category described above and (like those of Hogg [8] and Takeuchi) they are "quasi-linear" in that they are based on linear combinations of the order statistics with the optimal coefficients being estimated from the sample. The basic statistic from which the proposed estimators are developed consists of a linear combination of the order statistics with the coefficients constrained to be equal within each of several sets corresponding to blocks of successive order statistics. Thus, the number of distinct values assumed by the coefficients is some number k which is less than the sample size n, and one can hope to use the full sample to approximate the k optimal values for the coefficients. This type of statistic does not seem to have been considered before in the literature although it is similar in spirit to linear combinations of selected sample quantiles, and such statistics have been discussed

4

at length in many papers. The proposed estimation scheme and its asymptotic properties are detailed in Section 2. In Section 3 the special case $k = 2$ is discussed and some Monte Carlo results are given. The proof of the theorem stated in Section 2 is outlined in the Appendix.

2. The Proposed Estimator.

Let $X_1, X_2, \cdots, X_n$ represent independent identically distributed observations with common distribution function $F_\theta(x) = F(x-\theta)$ where F is a member of a family $\mathscr{F}$ of symmetric continuous distributions satisfying certain regularity conditions described below. Thus, θ represents the median of F_θ. Let $Y_1 \leq Y_2 \leq \cdots \leq Y_n$ represent the ordered X_i's (i.e., the "order statistics"). Suppose that n is even and let $T_{0n}, T_{1n}, \cdots, T_{kn}$ be a partition of the integers $1, 2, \cdots, n/2$, where T_{in} consists of t_{in} $\underline{successive}$ integers, so that $\sum_{i=0}^{k} t_{in} = n/2$. We first consider estimators of θ of the form
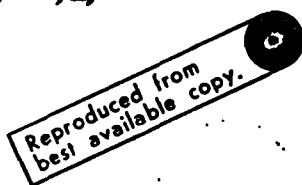
$$(1) \qquad \theta_{kn} = \frac{1}{n} \sum_{i=1}^{k} c_{in} S_{in} , \quad \text{where}$$

$$S_{in} = \sum_{j \in T_{in}} (Y_j + Y_{n-j+1}) , \quad i = 1, \cdots, k,$$

and the c_{in}'s are constants satisfying

$$(2) \qquad \sum_{i=1}^{k} c_{in} t_{in} = \frac{n}{2} .$$

Condition (2) insures that θ_{kn} is a location invariant estimator. Note that the t_{0n} smallest and largest order statistics have been "trimmed"

5

from the estimator θ_{kn} . This permits the development of substantially simpler formulae than would otherwise be possible. Suppose further that $c_{in} \to c_i$, $i = 1,2,\cdots,k$, and $t_{in}/n \to p_i$, $i = 0,1,2,\cdots,k$, as $n \to \infty$, where for each i, $p_i > 0$ and $\sum_{i=0}^{k} p_i = \frac{1}{2}$. Then we must also have from (2)

$$(3) \qquad \sum_{i=1}^{k} c_i p_i = cp' = \frac{1}{2} ,$$

where $c = (c_1, c_2, \cdots, c_k)$ and $p = (p_1, p_2, \cdots, p_k)$. Under these assumptions and the regularity conditions on F given below, for fixed k the sequence $\sqrt{n}\,(\theta_{kn} - \theta)$ is asymptotically normally distributed with mean zero and variance σ_k^2 given by

$$(4) \qquad \sigma_k^2 = 2 \sum_{i=0}^{k} \sum_{j=0}^{k} c_i c_j a_{ij} = 2cAc' ,$$

where $A = ((a_{ij}))_{k \times k}$ and the a_{ij}'s are defined as follows: Let ξ_i be the $(p_0 + p_1 + \cdots + p_i)^{th}$ quantile of F, i.e., $F(\xi_k) = p_0 + p_1 + \cdots p_{i-1}$ for $i = 1,2,\cdots,k-1$, and $\xi_k = 0$. Then for $i = 1,2,\cdots,k$

$$a_{ii} = 2\xi_i \int_{i-1}^{\xi_i} F(x)dx \;-\; 2\int_{\xi_{i-1}}^{\xi_i} xF(x)dx, \quad \text{and}$$

$$(5)$$

$$a_{ij} = a_{ji} = (\xi_j - \xi_{j-1}) \int_{\xi_{i-1}}^{\xi_i} F(x)dx, \quad i < j .$$

6

The asymptotic normality of θ_{kn} and the expression for σ_k^2 follow directly from Theorem 3, page 63, of Chernoff et al. [5].

At this point it is clear that since the a_{ij}'s are invariant under shifts, one could estimate them by replacing the ξ_i's by their sample analogs and $F(x)$ by the sample c.d.f. $F_n(x)$. One could then choose the c_i's so as to minimize (4) using the estimated a_{ij}'s and use the resulting c_i's in θ_{kn} . Such estimators would doubtless be (nearly) BAN for all $F \in \mathcal{F}$ for sufficiently large k and small p_0 (the trimming proportion). The computation of such estimates would, however, be quite complicated and the minimization process would require the inversion of a $k \times k$ matrix. We proceed, therefore, to introduce some simplifying approximations.

For $i = 1, 2, \cdots, k$ let

(6)
$$d_i = \xi_i - \xi_{i-1}, \quad \text{and} \quad b_i = \sum_{j=0}^{i-1} p_j + \frac{1}{2} p_i .$$

Let

(7)
$$B = \begin{pmatrix} b_1, b_1, \cdots, b_1 \\ b_1, b_2, \cdots, b_2 \\ \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \\ b_1, b_2, \cdots, b_k \end{pmatrix} ,$$

and

(8)
$$R = ((a_{ij}/d_i d_j))_{k \times k} - B ,$$

7

where the a_{ij}'s are given by (5).

Then if $\tilde{c} = (\tilde{c}_1, \tilde{c}_2, \cdots, \tilde{c}_k) = (c_1 d_1, c_2 d_2, \cdots, c_k d_k)$, and $\tilde{p} = (p_1/d_1, p_2/d_2, \cdots p_k/d_k)$, we have from (4), (7) and (8)

$$(9) \qquad \qquad \sigma_k^2 = \tilde{c}(B+R)\tilde{c}' \quad ,$$

and condition (3) becomes

$$(10) \qquad \qquad \tilde{c}\,\tilde{p}' = \frac{1}{2} \quad .$$

It is shown in the Appendix that for the cases of interest the contribution of $\tilde{c}R\tilde{c}'$ to σ_k^2 is negligible for large k . Hence we may minimize σ_k^2 approximately by minimizing $\tilde{c}B\tilde{c}'$ subject to condition (10). A matrix of the form (7) is non-singular if the b_i's are distinct, and a straightforward Lagrange multiplier argument establishes that the vector $\tilde{c}$ minimizing $\tilde{c}B\tilde{c}'$ subject to (10) is given by $\tilde{c} = \frac{1}{2}(\tilde{p}B^{-1}\tilde{p}')^{-1}B^{-1}\tilde{p}'$, and the corresponding approximate minimum of σ_k^2, denoted by $\tilde{\sigma}_k^2$, is given by $\tilde{\sigma}_k^2 = 2\tilde{c}B\tilde{c}' = \frac{1}{2}(\tilde{p}B^{-1}\tilde{p}')^{-1}$. Now letting $e = (e_1, e_2, \cdots, e_k)$ where $e_i = d_i^{-1} \times (i^{th}$ component of $B^{-1}\tilde{p}')$, we may write

$$(11) \qquad \qquad \tilde{\sigma}_k^2 = \frac{1}{2}\,(ep')^{-1},$$

and recalling that $c_i = \tilde{c}_i/d_i$, the (approximately) minimizing vector c is

$$(12) \qquad \qquad c = \tilde{\sigma}_k^2 e \quad .$$

8

The matrix B given by (7) is a Green's matrix and hence readily invertible yielding the Jacobi matrix $B^{-1} = ((b^{ij}))_{k \times k}$ with

$$b^{i,i+1} = b^{i+1,i} = \frac{-2}{p_i + p_{i+1}} \quad , \quad i = 1, 2, \cdots, k-1 \; ,$$

(13)
$$b^{11} = \frac{4p_0 + 4p_1 + 2p_2}{(2p_0 + p_1)(p_1 + p_2)} \quad ,$$

$$b^{ii} = \frac{2p_{i-1} + 4p_i + 2p_{i+1}}{(p_{i-1} + p_i)(p_i + p_{i+1})} \; , \quad i = 2, 3, \cdots, k-1$$

$$b^{kk} = \frac{2}{p_{k-1} + p_k} \quad ,$$

and all other $b^{ii} = 0$. Thus the components of the vector e are

$$e_1 = \left(\frac{2p_1(2p_0 + 2p_1 + p_2)}{(2p_0 + p_1)(p_1 + p_2)} \right) \frac{1}{d_1^2} - \left(\frac{2p_2}{p_1 + p_2} \right) \frac{1}{d_1 d_2} \quad ,$$

(14)
$$e_i = \frac{1}{d_i} \left[\left(\frac{2p_i(p_{i-1} + 2p_i + p_{i+1})}{(p_{i-1} + p_i)(p_i + p_{i+1})} \right) \frac{1}{d_i} - \left(\frac{2p_{i-1}}{p_{i-1} + p_i} \right) \frac{1}{d_{i-1}} - \left(\frac{2p_{i+1}}{p_i + p_{i+1}} \right) \frac{1}{d_{i+1}} \right],$$

$$i = 2, 3, \cdots, k-1,$$

$$e_k = \left(\frac{2p_k}{p_{k-1} + p_k} \right) \frac{1}{d_k^2} - \left(\frac{2p_{k-1}}{p_{k-1} + p_k} \right) \frac{1}{d_{k-1} d_k} \quad .$$

Note that for the special case $p_1 = p_2 = \cdots = p_k$, the formula for e_i for $i = 2, 3, \cdots, k-1$ simplifies to

(15)
$$e_i = \frac{1}{d_i} \left(\frac{2}{d_i} - \frac{1}{d_{i-1}} - \frac{1}{d_{i+1}} \right) \quad .$$

9

The only quantities appearing in (14) which depend on the distribution F are the d_i's which are differences of quantiles which may be effectively estimated by their sample analogs. For each i let $\hat{\xi}_i$ be the sample quantile corresponding to ξ_i and let $\hat{d}_i = \hat{\xi}_i - \hat{\xi}_{i-1}$. (More explicit formulas for the $\hat{d}_i$'s are offered in Section 3). Let $\hat{e}$ be the vector whose components $\hat{e}_i$, $i = 1,2,\cdots,k$, are obtained from (13) by replacing the d_i's by $\hat{d}_i$'s. Then following (11) and (12) let

$$(16) \qquad \hat{\sigma}^2_{kn} = \frac{1}{2}(\hat{e}p')^{-1} , \quad \text{and}$$

$$(17) \qquad \hat{c} = \hat{\sigma}^2_{kn}\hat{e} .$$

The proposed estimator of θ (by analogy with (1)) is

$$(18) \qquad \hat{\theta}_{kn} = \frac{1}{n} \sum_{i=1}^{k} \hat{c}_i S_{in} ,$$

where $\hat{c}_i$ is the ith component of $\hat{c}$.

It has been assumed so far that n is even. The case of n odd may be treated in various ways, perhaps the simplest of which is to delete the sample median from the order statistics and compute $\hat{c}$ on the basis of the even sample size $n-1$. The coefficient $\hat{c}_k$ should then be modified by multiplying by $t_{k,n-1}/(t_{k,n-1}+1)$ and the sample median should be added to $S_{k,n-1}$. Expression (18) may then be used with the $S_{i,n-1}$'s replacing the S_{in}'s.

To facilitate the discussion of the asymptotic behavior of $\hat{\theta}_{kn}$ we shall assume that for each k,

$$(19) \qquad p_1 = p_2 = \cdots = p_k = k^{-1}(\tfrac{1}{2} - p_0) = q_k \; ,$$

i.e., each S_{in} for $i \geq 1$ is a sum of (approximately) $2nq_k$ successive order statistics.

For any continuous, twice differentiable symmetric distribution $F(x)$ with density $f(x)$, the Fisher information for a location parameter when the np_0 smallest and np_0 largest order statistics have been deleted ("trimmed") from a sample of size n is given by

$$(20) \qquad I(F,p_0) = 2\int_{\xi_0}^{0} (f'(x))^2 (f(x))^{-1} dx + 2f^2(\xi_0)p_0^{-1} ,$$

where, as before, $F(\xi_0) = p_0$ (see, e.g. [5]). The untrimmed (full sample) case is obtained by setting $p_0 = 0$ ($\xi_0 = \infty$), and eliminating the second term on the right of (20). The Cramér-Rao lower bound for the variance of any estimator of location based on a trimmed sample of size n is then $(nI(F,p_0))^{-1}$.

We now define more precisely the family of distributions for which the proposed estimator is shown to be asymptotically (nearly) efficient.

Definition. Let $\mathcal{F}$ be a family of symmetric continuous distribution functions $F(x)$ with corresponding density functions $f(x) = F'(x)$ such that

(a) $f(x)$ and its first three derivatives exist and
are continuous with $f(x) > 0$ for all x, and
$xf(x) \to 0$ as $x \to \pm \infty$,

(b) $f(x)$ and its first three derivatives are uni-
formly bounded for all $F \epsilon \mathcal{F}$,

(c) $I(F,0)$ given by (20) is finite and uniformly
bounded away from zero for all $F \epsilon \mathcal{F}$, and

(d) for each $F \epsilon \mathcal{F}$ the conditions of Theorem 3 of
[5] are satisfied for estimators of the form (1).

The asymptotic behavior of the proposed estimator is summarized
in the following

Theorem: If $\hat{\theta}_{kn}$ and $\hat{\sigma}^2_{kn}$ are given by (18) and (16) respectively
and (19) is satisfied with the trimming proportion p_0 held fixed,
then for any $\epsilon > 0$ there exists a k_ϵ such that for each $k > k_\epsilon$,
for all $F \epsilon \mathcal{F}$, the sequence $\sqrt{n}(\hat{\theta}_{kn} - \theta_n)$ is asymptotically normally
distributed with mean zero and variance $\tau^2_k(F)$, as $n \to \infty$, where

$$(21) \qquad \tau^2_k(F) \leq \frac{1}{I(F,p_0)} + \epsilon$$

and furthermore, with probability approaching one as $n \to \infty$,

$$(22) \qquad |\hat{\sigma}^2_{kn} - \tau^2_k(F)| < \epsilon \quad .$$

The proof of the theorem is deferred to the Appendix.

If the trimming proportion p_0 is allowed to decrease we have the following immediate

Corollary: Under the conditions of the theorem, for any $\epsilon > 0$, there exist a k_ϵ and p_ϵ such that if the trimming proportion $p_0 = p_\epsilon$, then for each $k > k_\epsilon$, for all $F \epsilon \mathscr{F}$ the conclusions of the theorem hold with $I(F, p_0)$ replaced by $I(F, 0)$ in (21).

Thus, for a suitable choice of k and p_0 one can make $\hat{\theta}_{kn}$ as near to being BAN as desired, uniformly for all $F \epsilon \mathscr{F}$.

The actual family $\mathscr{F}$ for which the results of the theorem and its corollary hold is larger than that described in the above definition. It is shown in the Appendix that these results hold for the double exponential distribution (which is not in $\mathscr{F}$ as defined), and by implication, for other distributions having a finite number of simple discontinuities in f and f'.

One can conclude from the theorem that for any particular $F \epsilon \mathscr{F}$, there exist sequences $k_n \uparrow \infty$ and $p_{0n} \downarrow 0$, as $n \to \infty$, such that the corresponding sequence of estimators $\hat{\theta}_{k_n, n}$ (using trimming proportions p_{0n}) is BAN for F. Also $\hat{\sigma}^2_{k_n, n}$ is a consistent estimator of $(I(F, 0))^{-1}$. In order to assert the existence of a single pair of such sequences yielding a sequence of estimators which is BAN for all $F \epsilon \mathscr{F}$ it is necessary to further restrict $\mathscr{F}$ so as to insure uniform convergence to normality in the theorem cited from [5]. Questions concerning the rate behavior of the k_n and p_{0n} sequences necessary for this uniform BAN property can also be raised. These questions seem somewhat uninteresting in view of the results of the next section wherein it is shown that very small values of k suffice to produce good asymptotic efficiencies as

13

well as good small sample performance.

3. **The Cases** $k = 2$ **and** $k = 3$. Asymptotic efficiency-robustness is only a theoretical curiosity unless it is accompanied by satisfactory performance for small and moderate sample sizes. The estimator $\hat{\theta}_{kn}$ given by (18) reduces to the familiar trimmed mean for the case $k = 1$, so we concentrate on the next simplest cases where $k = 2$ or 3, with particular emphasis on $k = 2$.

It is convenient to introduce certain notational simplifications appropriate to the case of small k and n. Referring to the quantities involved in (1), let $r = t_{0n}$, $s = t_{1n} = t_{2n} = \cdots = t_{k-1,n}$, and $t = t_{kn} = \frac{n}{2} - r - (k-1)s$. Thus, for $k = 2$,

$$S_{0,n} = \sum_{j=1}^{r} (Y_j + Y_{n+1-j}) \ ,$$

(23)
$$S_{1,n} = \sum_{j=r+1}^{r+s} (Y_j + Y_{n+1-j}) \ ,$$

$$S_{2,n} = \sum_{j=r+s+1}^{n/2} (Y_j + Y_{n+1-j}) \ .$$

We now present explicit formulas for the estimates $\hat{d}_i$ of $d_i = \xi_i - \xi_{i-1}$. Any suitably defined sample analogs of ξ_i and ξ_{i-1} will suffice to define $\hat{d}_i$ but since ξ_i represents the quantile which is the upper endpoint of the quantile range corresponding to the sample quantiles contributing to S_{in} and the lower endpoint of the corresponding range for $S_{i+1,n}$ it seems reasonable to choose for $\hat{\xi}_i$ the midpoint between the largest order statistic contributing to S_{in} and

14

the smallest contributing to $S_{i+1,n}$. This leads to estimates d_i given by

$$\hat{d}_i = \frac{1}{4}(Y_{r+is} + Y_{r+is+1} + Y_{n-r-(i-1)s} + Y_{n-r-(i-1)s+1}$$

(24)
$$- Y_{r+(i-1)s} - Y_{r+(i-1)s+1} - Y_{n-r-is} - Y_{n-r-is+1}), \quad i=1,2,\ldots,k-1,$$

$$\hat{d}_k = \frac{1}{4}\left(Y_{(n/2)+t} + Y_{(n/2)+t+1} - Y_{(n/2)-t} - Y_{(n/2)-t+1} \right).$$

Thus, for $k = 2$ and even n, noting that $p_0 = r/n$, $p_1 = s/n$, and $p_2 = t/n$, we have from (14) and the definition of $\hat{e}$,

(25)
$$\hat{e}_1 = \left(\frac{2s(2r+2s+t)}{(2r+s)(s+t)}\right) \frac{1}{\hat{d}_1^2} - \left(\frac{2t}{s+t}\right) \frac{1}{\hat{d}_1\hat{d}_2},$$

$$\hat{e}_2 = \left(\frac{2t}{s+t}\right) \frac{1}{\hat{d}_2^2} - \left(\frac{2s}{s+t}\right) \frac{1}{\hat{d}_1\hat{d}_2},$$

and from (16), (17), and (18),

(26)
$$\hat{\sigma}_{2,n}^2 = \frac{n}{2(s\hat{e}_1+t\hat{e}_2)}, \quad \text{and}$$

(27)
$$\hat{\theta}_{2,n} = \frac{\hat{e}_1 S_{1,n} + \hat{e}_2 S_{2,n}}{2(s\hat{e}_1+t\hat{e}_2)}.$$

Similar simple formulas result for the case $k = 3$.

For given values of p_0, p_1, and p_2 and any particular distribution F, the asymptotic variance of $\hat{\theta}_{2,n}$ may be calculated by first calculating the four a_{ij}'s given by (5) and $c = (c_1, c_2)$ given by (12)

15

and then applying (4). Asymptotic variances for five cases were computed and the efficiencies relative to $(I(F,0))^{-1}$ are shown below.

Table 1

Asymptotic Efficiencies for $\hat{\theta}_{2,n}$ when $p_0 = .05$, $p_1 = p_2 = .225$

Distribution	Normal	C.N.(.10)	Cauchy	Dble. Exp.	Logistic
Efficiency (%)	94.4	93.6	97.2	84.5	97.5

A Monte Carlo experiment was performed for the cases $k = 2$ and 3, using 4000 replications, for the normal, contaminated normal (C.N.(.10), Cauchy, double exponential, and logistic distributions. (The C.N.(.10) distribution is a standard normal with a 10% contamination by a normal with mean zero and variance $= 9$.) Sample sizes $n = 10, 20, 40$, and 80 were used, and the trimming proportion p_0 (from each tail) was taken to be .05 for all cases except $n = 10$ for which $p_0 = .10$. For the case $k = 2$, for $n > 10$, the quantities s and t were chosen in three ways: (i) $2s = t$, (ii) $s = 2t$, and (iii) $s \approx t$ (subject to the limitation that s and t must be positive integers with $s + t = \frac{n}{2} - r$). For $n = 10$, s and t were set equal to two. The case $k = 3$ was computed for $n = 20, 40$, and 80 with $p_0 = .05$ and $s = t$.

Table 2 shows the mean squared errors (MSE) and relative efficiencies of $\hat{\theta}_{2,n}$, and the means of $\hat{\sigma}^2_{2,n}$, for the case $k = 2$, $s \approx t$. Ideally the small sample efficiencies should be computed relative to the variances of the minimum variance location invariant (Pitman) estimators, but these were only available for the double exponential distribution (from

16

Table 2

Efficiencies for Case $k = 2$, $s \approx t$

Sample Size	Distribution	Var(BLUE)	MSE($\hat{\theta}_{2,n}$)	$\hat{\theta}^2_{2,n}/n$	Eff.($\theta_{2,n}$)
$n = 10$	Normal	.1000	.1184	.0892	84%
	C.N.(.10)	.1358	.1488	.1245	91
	Cauchy	.3261	.4229	.4841	77
	Dble. Exp.	.1399	.1645	.1399	85
	Logistic	.3073	.3446	.2750	89
$n = 20$	Normal	.0500	.0579	.0525	86
	C.N.(.10)	.0657	.0710	.0712	93
	Cauchy	.1256	.1304	.1429	96
	Dble. Exp.	.0637	.0703	.0707	91
	Logistic	.1520	.1668	.1573	91
$n = 40$	Normal	.0250	.0284	.0281	88
	C.N.(.10)	.0315*	.0354	.0364	89
	Cauchy	.0500*	.0589	.0646	85
	Dble. Exp.	.0297**	.0333	.0353	89
	Logistic	.0750*	.0817	.0822	92
$n = 80$	Normal	.0125	.0137	.0147	91
	C.N.(.10)	.0157*	.0170	.0186	92
	Cauchy	.0250*	.0269	.0300	93
	Dble. Exp.	.0142**	.0155	.0174	92
	Logistic	.0375	.0393	.0424	95

* Cramér-Rao bound **Pitman estimator

Table 3

Confidence Interval Error Probabilities

$$P\{|\hat{\theta}_n| > \gamma\hat{\theta}_{kn}/\sqrt{n}\} \ , \ (k=2, \ s \approx t)$$

Distribution	Sample Size(n)	$\gamma = 1.645$	$\gamma = 1.960$	$\gamma = 2.576$
Normal	10	.208	.143	.074
	20	.145	.087	.039
	40	.119	.066	.020
	80	.095	.048	.011
C.N.(.10)	10	.196	.136	.068
	20	.135	.083	.036
	40	.117	.065	.021
	80	.094	.050	.011
Cauchy	10	.148	.100	.047
	20	.124	.076	.030
	40	.112	.057	.017
	80	.095	.048	.011
Dble. Exp.	10	.176	.120	.058
	20	.128	.074	.028
	40	.111	.061	.015
	80	.092	.044	.090
Logistic	10	.200	.138	.070
	20	.140	.088	.038
	40	.117	.065	.019
	80	.094	.048	.011
"t"-Dist.	10	.135	.082	.030
$P\{|t_{n-1}| > \gamma\}$	20	.117	.065	.019
	40	.108	.057	.014
	80	.104	.054	.012
Normal $P\{|\bar{X}_n| > \frac{\gamma}{\sqrt{n}}\}$	all n	.100	.050	.010

18

Table 4

"Best Case" Efficiencies (k = 2)

	Sample Size					
Distribution	n=20		n=40		n=80	
	Best(s,t)	Eff.($\hat{\theta}_{2,n}$)	Best(s,t)	Eff.($\hat{\theta}_{2,n}$)	Best(s,t)	Eff.($\hat{\theta}_{2,n}$)
Normal	(3,6)	90%	(6,12)	90%	(12,24)	94%
C.N.(.10)	(3,6)	95	(6,12)	92	(12,24)	94
Cauchy	(6,3)	100	(12,6)	87	(24,12)	96
Dble. Exp.	(6,3)	92	(12,6)	93	(24,12)	97
Logistic	(4,5)	92	(9,9)	92	(12,24)	96

Table 5

Ratio of Variance for k = 3 to Variance for k = 2

for the Case s ≈ t

Distribution	n=20	n=40	n=80
Normal	1.03	1.07	1.05
C.N.(.10)	1.04	1.05	1.05
Cauchy	1.19	1.11	1.08
Dble. Exp.	1.09	1.08	1.06
Logistic	1.03	1.07	1.06

19

V. Miké [10]). For the other distributions the variances of the BLUE's were used when available, and for the larger sample sizes the Cramér-Rao bounds were used which accounts for the apparent slight drop in several of the efficiencies at $n=40$ compared to those at $n=20$. The fact that the efficiencies shown for the double exponential case are relatively high compared to the asymptotic value given in Table 1 is a consequence of the rather slow rate at which the variance of the Pitman estimator approaches its asymptotic value as the sample size increases.

Table 3 shows the estimated error (non-coverage) probabilities for confidence intervals of the form $\hat{\theta}_{2,n} \pm \gamma n^{-1/2} \hat{\sigma}_{2,n}$. The values of γ were taken (arbitrarily) to be percentage points of the standard normal distribution. For comparison purposes the error probabilities are given for each n for the corresponding confidence intervals based on the t-distribution (assuming normally distributed observations). The distribution of the "studentized" version of $\hat{\theta}_{2,n}$ shows a somewhat greater spread than the corresponding t-distribution for the smaller sample sizes. For each pair of values of n and γ the error probabilities are rather similar for all five of the distributions considered, and they do not differ much from those for the corresponding t-distribution for $n=40$ and 80.

Table 4 shows for each distribution the most favorable pair of values of s and t (among the three possibilities tested for each n) together with the corresponding efficiencies for $n = 20$, 40 and 80. For all

three sample sizes the large values of t, leading to large groups of central order statistics receiving weight $\hat{c}_2$, tend to be favorable for the relatively short-tailed distributions (normal, contaminated normal, and logistic), while the reverse is true for the long-tailed cases (Cauchy and double exponential). This provides considerable justification for the choice $s = t$ as a suitable compromise.

Table 5 shows the ratio of the variance of $\hat{\theta}_{3,n}$ to the variance of $\hat{\theta}_{2,n}$ when $s \approx t$, for $n = 20, 40$, and 80, for each of the five distributions. The uniform superiority of the estimator using $k = 2$ must be attributed to the additional sampling fluctuation introduced when an additional coefficient is estimated in the case $k = 3$, together with the fact that the efficiencies are already high for the case $k = 2$. In a preliminary Monte Carlo investigation, larger values of k were considered and in all cases the variances of the estimators for a given sample size tended to increase with increasing k .

In interpreting the Monte Carlo results given above it should be borne in mind that the standard deviation associated with each of the efficiencies is on the order of two percentage points. It is perhaps worth mentioning that the ordinary congruence-type pseudo-random numbers proved to be completely inadequate for this investigation, and it was found necessary to introduce a suitable "re-randomization" procedure.

The possibility exists that the performance of the estimator $\hat{\theta}_{2,n}$ could be slightly improved by a minor modification as indicated in the following

Remark: The quantities $\hat{e}_1$ and $\hat{e}_2$ appearing in the definition (27)

21

of $\hat{\theta}_{2,n}$ are of the form

$$\hat{e}_1 = \alpha_1 \hat{d}_1^{-2} - \alpha_2 (\hat{d}_1 \hat{d}_2)^{-1} ,$$

$$\hat{e}_2 = \beta_1 \hat{d}_2^{-2} - \beta_2 (\hat{d}_1 \hat{d}_2)^{-1} ,$$

where $\alpha_i, \beta_i, i = 1, 2,$ are determined by (25). Since the formulae given in (25) were obtained by an asymptotic argument valid for large k, it is remarkable that they work well for $k = 2$. It seems entirely possible that a somewhat different choice for the α_i's and β_i's could result in uniformly improved performance. The author conjectures that a slight modification of α_1 (only) could lead to improved efficiency for short-tailed distributions like the normal without materially impairing the performance of the estimator for long-tailed distributions.

Appendix

To prove the theorem of Section 2 we first show that $\sqrt{n}\,(\theta_{kn}-\theta)$ is asymptotically normal with mean zero and variance

$$(28) \qquad \sigma_k^2 \leq \frac{1}{I(F,p_0)} + \epsilon \; ,$$

for arbitrary fixed $\epsilon > 0$, for each sufficiently large k, for all $F \in \mathcal{F}$, when θ_{kn} is given by (1) with the c_i's given by (12). The validity of (20) follows from conditions (a) and (c) of the definition of $\mathcal{F}$, and asymptotic normality follows from condition (d). By virtue of expression (9) it therefore suffices to show (i) that (28) holds with σ_k^2 replaced by $\tilde{\sigma}_k^2$ given by (11), and (ii) that $|\tilde{c}\,R\,\tilde{c}'| < \epsilon$ for all sufficiently large k, for all $F \in \mathcal{F}$.

To verify assertion (i), let $F^{-1}(u) = G(u)$, and referring to (19) let $r_i = p_0 + (i-1)q_k$. Then by definition (6) of d_i we may write

$$d_i = G(r_i + q_k) - G(r_i) \; , \quad i = 1, 2, \ldots, k,$$

$$d_{i+1} = G(r_i + 2q_k) - G(r_i + q_k) \; , \quad i = 1, 2, \ldots, k-1,$$

$$d_{i-1} = G(r_i) - G(r_i - q_k), \quad i = 2, 3, \ldots, k.$$

Now let G_i', G_i'' and G_i''' represent the first three derivatives of G evaluated at r_i. Then under conditions (a) and (b) of the definition of $\mathcal{F}$, a straightforward Taylor's series expansion in terms of q_k yields

$$d_i^{-1} = q_k^{-1}(G_i')^{-1} - \frac{1}{2}(G_i')^{-2}G_i'' + q_k[\frac{1}{4}(G_i'')^2(G_i')^{-3} - \frac{1}{6}G_i'''(G_i')^{-2}] + \zeta q_k^2 ,$$

$$d_{i+1}^{-1} = q_k^{-1}(G_i')^{-1} - \frac{3}{2}(G_i')^{-2}G_i'' + q_k[\frac{9}{4}(G_i'')^2(G_i')^{-3} - \frac{7}{6}G_i'''(G_i')^{-2}] + \zeta q_k^2 ,$$

$$d_{i-1}^{-1} = q_k(G_i')^{-1} + \frac{1}{2}(G_i')^{-2}G_i'' + q_k[\frac{1}{4}(G_i'')^2(G_i')^{-3} - \frac{1}{6}G_i'''(G_i')^{-2}] + \zeta q_k^2 ,$$

where ζ represents a generic uniformly bounded function of i, k, and G. Hence by (15), for $i = 2, 3, \ldots, k-1$,

$$e_i = - [2(G_i'')^2(G_i')^{-4} - G_i'''(G_i')^{-3}] + \zeta q_k .$$

Letting f_i, f_i', and f_i'' be f and its corresponding derivatives evaluated at $G(r_i)$ we have

$$G_i' = (f_i)^{-1}, \quad G_i'' = -f_i'(f_i)^{-3}, \quad G_i''' = 3(f_i')(f_i)^{-5} - f_i''(f_i)^{-4} ,$$

so that $e_i = -[f_i''(f_i)^{-1} - (f_i')^2(f_i)^{-2}] + \zeta q_k$.

Also, writing

$$d_k = G(\tfrac{1}{2}) - G(\tfrac{1}{2} - q_k),$$

$$d_{k-1} = G(\tfrac{1}{2} - q_k) - G(\tfrac{1}{2} - 2q_k)$$

and referring to (14), we obtain by a similar calculation

$$(29) \qquad e_k = G'''(\tfrac{1}{2})[G'(\tfrac{1}{2})]^{-3} + \zeta q_k = - \frac{f''(0)}{f(0)} + \zeta q_k .$$

Hence, as $k \to \infty$,

$$(30) \qquad q_k \sum_{i=2}^{k} e_i \to - \int_{p_0}^{\frac{1}{2}} \left\{ \frac{f''(G(u))}{f(G(u))} - \frac{[f'(G(u))]^2}{f^2(G(u))} \right\} du ,$$

$$= - \int_{G(p_0)}^{0} \left\{ \frac{\partial^2}{\partial x^2} \log f(x) \right\} f(x) dx = \int_{G(p_0)}^{0} (f'(x))^2 (f(x))^{-1} dx + f_1' ,$$

and the convergence is uniform for all $F \in \mathcal{F}$. Again referring to (14), we obtain

$$(31) \qquad e_1 = q_k^{-1}(\frac{1}{p_0} f_1^2 - f_1') + \zeta .$$

Thus, from (11), (19), (20), (30) and (31) we have

$$\frac{1}{\tilde{\sigma}_k^2} = 2q_k \sum_{i=1}^{k} e_i \to I(F, p_0)$$

as $k \to \infty$, uniformly for all $F \in \mathcal{F}$. This together with condition (c) of the definition of $\mathcal{F}$ implies the truth of assertion (i).

To prove assertion (ii), let the elements of the matrix R given by (8) be denoted by r_{ij}, $i, j = 1, 2, \ldots, k$. Then by (5) and (6) we have for $i < j$

$$r_{ij} = r_{ji} = d_i^{-1} \int_{G(r_i)}^{G(r_i + q_k)} f(x) dx - b_i ,$$

and

$$r_{ii} = 2d_i^{-2} \left\{ G(r_i+q_k) \int_{G(r_i)}^{G(r_i+q_k)} F(x)dx - \int_{G(r_i)}^{G(r_i+q_k)} x\, F(x)dx \right\} - b_i \ ,$$

where by (6) and (19), $b_i = r_i + \frac{1}{2} q_k$.

Expansion in Taylor's series with respect to q_k yields for $i < j$,

$$r_{ij} = r_i + \frac{1}{2} q_k G_i' f_i + \zeta q_i^2 - b_i = \zeta q_k^2 \ ,$$

and

$$r_{ii} = r_i + q_k G_i' f_i - b_i + \zeta q_k^2 = \frac{1}{2} q_k + \zeta q_k^2 \ .$$

Since $q_k = k^{-1}(\frac{1}{2} - p_0)$ we see that $r_{ij} = O(k^{-2})$ uniformly in i,j and $F \in \mathcal{F}$ for $i \neq j$, and $r_{ii} = O(k^{-1})$ uniformly in i and $F \in \mathcal{F}$. Furthermore, since $\tilde{c}_i = c_i d_i$ where the c_i's are given by (12) we see that $\tilde{c}_i = O(k^{-1})$ uniformly in i and $F \in \mathcal{F}$, so that $\tilde{c} R \tilde{c}' = O(k^{-2})$, uniformly in $F \in \mathcal{F}$ which proves assertion (ii).

Conclusion (21) of the theorem follows from (28) provided that for any fixed k, $\sqrt{n} (\hat{\theta}_{kn} - \theta_{kn}) \to 0$, in probability, as $n \to \infty$, since then $\tau_k^2(F) = \sigma_k^2$. But from (18) we have

$$\sqrt{n} (\hat{\theta}_{kn} - \theta_{kn}) = \sum_{i=1}^{k} (\hat{c}_i - c_i)(\frac{1}{\sqrt{n}} S_{in}) \ .$$

Now the $\hat{c}_i$'s given by (17) are consistent estimates of the corresponding c_i's given by (12) since the $\hat{d}_i$'s are consistent estimates of the corresponding d_i's. The desired result follows from the

26

boundedness, in probability, of the quantities $\frac{1}{\sqrt{n}} S_{in}$, as $n \to \infty$.
Conclusion (22) now follows from the fact that $\hat{\sigma}_k^2$ is a consistent estimate
of $\tilde{\sigma}_k^2$ which by assertion (11) is uniformly close to σ_k^2 for sufficiently
large k. This completes the proof of the theorem.

We consider now the double exponential distribution with density
$f(x) = \frac{1}{2} e^{-|x|}$. The expressions for e_i obtained above are valid for
$i=1,2,\ldots,k-1$. From (31) we see that $e_1 = \zeta$ for this case. Also
$\frac{\partial^2}{\partial x^2} \log f(x) = 0$, for $x < 0$, so that $q_k \sum_{i=1}^{k-1} e_i \to 0$, as $k \to \infty$.
Expression (29) for e_k is not valid for this case since $f'(x)$ is
not continuous at $x = 0$. Direct calculation shows that $e_k = 4q_k^{-1}+0(1)$.
Therefore, in the limit, the only non-zero c_i is c_k, so that, as k
increases, θ_{kn} is essentially an average of a decreasing proportion of
central order statistics. The conclusions of the theorem are clearly
correct for the corresponding sequence $\hat{\theta}_{kn}$ even though the double
exponential distribution is not in the family $\mathcal{F}$ as defined.

REFERENCES

[1] Bhattacharya, P.K., "Efficient Estimation of a Shift Parameter from Grouped Data," Annals of Mathematical Statistics, 38 (1967), 1770-87.

[2] Birnbaum, A. and Laska, E., "Optimal Robustness: A General Method, with Applications to Linear Estimators of Location," Journal of the American Statistical Association, 62 (1967), 1230-40.

[3] ____, and Miké, V., "Asymptotically Robust Estimators of Location," Journal of the American Statistical Association, 65 (1970), 1265-82.

[4] ____, Laska, E. and Meisner, M., "Optimally Robust Linear Estimators of Location," Journal of the American Statistical Association, (in press).

[5] Chernoff, H., Gastwirth, J.L. and Johns, M.V., Jr., "Asymptotic Distribution of Linear Combinations of Order Statistics," Annals of Mathematical Statistics, 31 (1967), 52-72.

[6] Gastwirth, J.L., and Rubin, H., "On Robust Linear Estimators," Annals of Mathematical Statistics, 40 (1969), 24-39.

[7] ____, and Cohen, M.L., "Small Sample Behavior of Some Robust Linear Estimators of Location," Journal of the American Statistical Association, 65 (1970), 946-73.

[8]   Hogg, R.V., "Some Observations on Robust Estimation," <u>Journal of</u>
      <u>the American Statistical Association</u>, 62 (1967), 1179-86.

[9]   Jaeckel, L.A., "Robust Estimates of Location," University of
      California Ph.D. dissertation (1969), unpublished.

[10] Miké, V., "Efficiency-Robust Systematic Linear Estimators of
      Location," <u>Journal of the American Statistical Association</u>,
      (in press).

[11] Stein, C., "Efficient Nonparametric Testing and Estimation,"
      <u>Proceedings of the Third Berkeley Symposium on Mathematical</u>
      <u>Statistics and Probability</u>, 7 (1956), 187-96, University of
      California Press.

[12] Switzer, P., "Efficiency Robustness," <u>Proceedings of the Sixth</u>
      <u>Berkeley Symposium on Mathematical Statistics and Probability</u>,
      University of California Press, (in press).

[13] Takeuchi, K., "A Uniformly Asymptotically Efficient Robust
      Estimator of a Location Parameter," <u>Journal of the American</u>
      <u>Statistical Association</u>, (in press).

[14] Van Eeden, C., "Efficiency-Robust Estimation of Location,"
      <u>Annals of Mathematical Statistics</u>, 41 (1970) 172-81.